

Performance of adapting non-native speech in isolated speech recognizer

ABSTRACT

This paper presents the pronunciation performance between native and non-native speakers of Malay sounds in isolated speech recognizer (ISR). Speaker adaptation methods are combined to solve the performance decrease that recognizers are faced with native and non-native speech of speaker-independent (SI) models. Often, speech recognition performance degrades drastically if the recognizer which has been trained with native speech is exposed to non-native speech even though speech recognition systems have reached a certain level of maturity. Two experiments was performed to show the recognition accuracy of the baseline models trained with native dataset was drastically low for the non-native speakers from non-Malay group than for the native ones. Acoustic deviation has been discovered as one of important factors affecting the performance of the ISR. In this experiment, an acoustic technique has been implemented to compare the performance on native and non-native speech. We explore how acoustic models can be adapted to better recognize the non-native speech. The experiments show that there are many problems arise such as adaptation methods and the non-native pronunciation pattern that remains to be investigated. In future, it will be necessary to improve speaker adaptation methods by incorporating more extensive knowledge of speaker variation at both the acoustic and the pronunciation level.

Keyword: Isolated speech; Native speaker; Non-native speaker; Maximum Likelihood Linear Regression (MLLR); Maximum a Posteriori (MAP)